INTERNATIONAL JOURNAL OF PHARMACEUTICAL SCIENCES
[ISSN: 0975-4725; CODEN(USA): IJPS00]
Journal Homepage: https://www.ijpsjournal.com

Review Article

AI-Driven Predictive Modelling for Stability Assessment of Biologic Therapeutics

Ankita Gidde*

LSHGCT's Gahlot Institute of Pharmacy, Koparkhairane Sector 14, Navi Mumbai, Maharashtra, India, 400709

ARTICLE INFO

Published: 3 Jun 2026
Keywords:
Machine Learning, Deep Learning, Drug Stability, AlphaFold, biologic Therapeutics, Predictive Analytics.
DOI:
10.5281/zenodo.20532275

ABSTRACT

Biologic products, including proteins, monoclonal antibodies, vaccines, and nucleic-acid–based therapeutics, represent an advanced class of pharmaceuticals characterized by structural complexity and sensitivity to environmental conditions. Ensuring the stability of these biomolecules is critical for maintaining product safety, efficacy, and overall quality throughout their lifecycle. Conventional stability testing approaches are often time-consuming and experimentally intensive, creating a need for innovative predictive strategies. Artificial intelligence (AI) has emerged as a transformative tool capable of reshaping stability assessment and development of biologic products. AI techniques, particularly machine learning (ML) and deep learning (DL), enable the integration and analysis of large multidimensional datasets to identify degradation patterns, predict shelf life, and support formulation optimization. Advanced computational models, including structure prediction platforms such as AlphaFold and de novo design methodologies, facilitate improved understanding of protein folding, molecular interactions, and conformational stability, thereby supporting rational biologic development.AI-driven models further evaluate critical factors influencing biologic stability, including formulation composition, excipient compatibility, processing conditions, and environmental variables such as temperature, pH, light exposure, and storage stress. These predictive capabilities allow early identification of stability risks and enable data-guided decision-making during product development and manufacturing. Overall, AI demonstrates significant potential to transform biologic stability studies by accelerating development timelines, reducing experimental burden, and improving predictive accuracy. The integration of AI-based analytical frameworks into biologic development pipelines is expected to enhance product quality assurance and support future regulatory and pharmaceutical innovation.

INTRODUCTION

Over the past three decades, the therapeutic landscape has undergone a significant

***Corresponding Author:** Ankita Gidde
Address: *LSHGCT's Gahlot Institute of Pharmacy, Koparkhairane Sector 14, Navi Mumbai, Maharashtra, India, 400709*
Email : ankitagidde06@gmail.com

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

transformation with the rapid emergence of biologic drugs as a major component of modern medicine. Biologics are complex therapeutic agents primarily composed of proteins and nucleic-acid–based molecules, including monoclonal antibodies, therapeutic enzymes, messenger RNA (mRNA), and small interfering RNA (siRNA). Unlike conventional small-molecule drugs that typically act through simple receptor or enzyme interactions, biologics exhibit highly specific and multifunctional mechanisms, enabling targeted modulation of biological pathways. This specificity has expanded treatment possibilities for complex diseases such as cancer, autoimmune disorders, genetic diseases, and infectious conditions where traditional therapies often show limited effectiveness [3]. Despite their therapeutic advantages, biologics present substantial development challenges due to their large molecular size, structural complexity, and sensitivity to physicochemical conditions. The immense diversity of possible protein sequences and conformations makes experimental exploration labour-intensive, costly, and time-consuming. However, advances in high-throughput biological data generation combined with the rapid evolution of artificial intelligence (AI), particularly deep learning technologies, are reshaping biologic research. AI-driven computational models can analyse complex biological datasets, predict structure–function relationships, and support rational design strategies, thereby transitioning biologic development from empirical experimentation toward data-driven innovation [1].

Artificial intelligence (AI) has evolved from a theoretical concept into a transformative technology impacting multiple scientific and industrial fields. Advances in computational power, data availability, and machine learning have accelerated its adoption in healthcare and pharmaceutical sciences.AI refers to computer-based systems capable of performing tasks that require human intelligence, such as learning from data, pattern recognition, prediction, and decision-making. Modern AI systems improve performance through adaptive learning while operating under defined regulatory frameworks. In pharmaceutical research and development, AI is widely applied in drug discovery, diagnostics, manufacturing optimization, and therapeutic development, enabling efficient analysis of complex biological data and supporting improved biologic drug design and stability assessment [2]. Common Types of Biologic Therapeutics :-Biologic therapeutics include diverse advanced medicinal products such as monoclonal antibodies, recombinant proteins, vaccines, and gene therapies. These biologics provide targeted treatment for cancer, metabolic, infectious, and genetic diseases but possess complex structures requiring advanced analytical characterization. Ensuring stability, potency, and product quality across these modalities highlights the growing need for sophisticated analytical and predictive evaluation strategies [5]. Artificial intelligence (AI) and predictive analytics are transforming drug stability assessment by addressing the limitations of conventional experimental approaches. Machine learning and deep learning models enable rapid prediction of long-term stability, simulation of environmental stress conditions, and optimization of formulation parameters influencing product quality. By integrating large datasets with real-time monitoring, AI supports continuous process verification, data-driven decision-making, and improved stability management of complex biologic therapeutics [2]. This review provides a structured overview of transformative application of artificial intelligence in biologics, covering fundamental AI methodologies and their roles in sequence and structure prediction and generative modelling [1].

CHALLENGES IN STABILITY OF BIOLOGICS

Biologics are unstable compared to small molecule drugs because they possess complex three-dimensional structure. Their stability is affected by chemical, physical, and environment factors. Chemical instability such as deamidation involves the conversion of asparagine (Asn) or glutamine (Gln) residues into acidic forms, leading to structural alteration and reduced biological activity. Deamidation is challenge in monoclonal antibodies, where it may occur in complementarity-determining regions, leading to reduced target affinity [17]. Oxidation commonly affects methionine, cysteine, tryptophan, tyrosine residues and it is induced by oxygen, light exposure or trace metal ions, resulting in loss of potency and aggregation. Hydrolysis results from peptide bond cleavage under extreme pH or moisture conditions, causing fragmentation of protein molecules. Disulfide bond instability arises from formation or scrambling of disulfide bridges, leading to protein misfolding and decreased stability [18]. In the physical instability include protein denaturation, aggregation, precipitation and surface adsorption that affect higher-order structure without changing chemical composition. Environmental conditions such as pH variations, temperature stress, light exposure, and mechanical agitation further accelerate degradation pathways. Additionally, manufacturing and storage processes including lyophilization, and transportation can compromise product stability. Collectively, these factors may lead to loss of potency, reduced efficacy, increased immunogenicity, and shortened shelf life [19].

LIMITATION OF CONVENTIONAL STABILITY ASSESSMENT OF BIOLOGICS

Conventional stability testing methods form the foundation of pharmaceutical quality evaluation and regulatory assessment. However, despite their widespread application, these approaches exhibit several scientific and practical limitations, particularly for complex biologic products. Traditional stability studies are primarily conducted under in-vitro experimental conditions using purified proteins in dilute buffer systems. These simplified environments fail to reproduce physiological conditions such as molecular crowding, cellular interactions, and biological complexity. As a result, experimentally determined stability may not accurately reflect real in-vivo protein behaviour.

Real-time stability studies often require 12–36 months to generate data [13]. Such extended timelines delay formulation optimization, scale-up activities, and product launch. Although accelerated stability studies are used to reduce time, elevated temperature and humidity conditions may induce degradation pathways different from those occurring under actual storage conditions, reducing predictive reliability [7]. Conventional approaches are largely reactive rather than predictive. Stability failures are typically detected only after significant degradation has occurred, limiting opportunities for early intervention and proactive formulation design. Biologic molecules exhibit dynamic conformational behaviour and may exist in multiple structural states depending on their environment or functional location. Conventional methods often assume a single stable native structure, thereby oversimplifying protein stability mechanisms. Stability evaluation is usually performed on isolated and purified proteins, excluding cofactors, binding partners, and intracellular conditions that significantly influence protein folding, aggregation, and degradation pathways. This reduces biological relevance[15]. Prediction of long-term stability frequently relies on extrapolation of accelerated stability data. Such

extrapolation introduces uncertainty, particularly for complex biologics, combination products, and modified-release formulations. Traditional statistical tools and regression analyses generally assume linear degradation behaviour. However, biologic stability depends on nonlinear interactions among formulation components, environmental factors, and process variables, which conventional models often fail to capture [21]. Modern biologic formulations generate large and complex datasets. Conventional analytical and statistical approaches show restricted capability in analysing multidimensional data, reducing sensitivity in detecting subtle degradation trends. Conventional stability programs require extensive laboratory infrastructure, stability chambers, analytical testing, repeated sampling, and skilled personnel. These requirements significantly increase operational costs and workload. Large volumes of stability data are commonly processed manually, increasing the risk of data management errors, inefficiencies, and variability in interpretation. Emerging pharmaceutical paradigms such as continuous manufacturing, real-time release testing, and lifecycle-based quality management demand rapid, data-driven decision making. Conventional stability approaches are insufficient to support these modern requirements. These limitations highlight the need for advanced predictive and data driven stability assessment strategies. Artificial intelligence and machine learning approaches have emerged as powerful tools capable of mode ling complex degradation behaviour, enabling proactive stability prediction and improved quality management for biologic products [21,22].

FUNDAMENTAL OF ARTIFICIAL INTELLIGENCE IN BIOLOGICS -

Machine Learning (ML) in Biologics

Traditional machine learning approaches have played a significant role in biologics research by relying on manually designed features such as amino acid composition, hydrophobicity, and structural motifs. Algorithms like Support Vector Machines (SVM), Random Forest (RF), and Multi-Layer Perceptron (MLP) have been widely applied to predict antigen–antibody interactions, assess immunogenicity, and model pharmacokinetic properties of biologics. These models are particularly effective in classification and regression tasks, enabling researchers to group biologics with similar structural or functional properties and to identify promising therapeutic candidates. However, the dependence on handcrafted descriptors often limits their ability to capture hidden biological patterns, making feature engineering both time-consuming and prone to bias. Therefore, while not fully capturing the complexity of biological systems, traditional ML methods still form an essential foundation in the development of AI-driven biologics research pipelines [23,24].

Deep Learning (DL) in Biologics

Deep learning approaches have emerged as a powerful alternative by automatically extracting complex features from raw biological data such as protein sequences, three-dimensional structures, and omics datasets. Techniques including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Autoencoders, Deep Neural Networks (DNNs), and Restricted Boltzmann Networks (RBNs) have been successfully applied to biologics research. These models excel at multi-layer feature extraction, transforming simple descriptors into higher-order representations that improve predictive accuracy while reducing generalization errors. Deep learning has enabled applications such as designing stable therapeutic proteins,

identifying epitopes for vaccine development, and predicting off-target effects or toxicity. Unlike traditional ML, DL models scale effectively with large datasets and can leverage transfer learning, where pretrained models on extensive protein databases are fine-tuned for specific biologic tasks. Despite challenges in interpretability, deep learning offers a transformative advantage in biologics by uncovering complex biological relationships that were previously inaccessible through manual feature engineering [24].

AlphaFold in Predictive Modelling for Biologic Stability

AlphaFold, developed by DeepMind, represents a major breakthrough in AI-based protein structure prediction. It employs advanced neural network architectures that integrate evolutionary, physical, and geometric constraints of proteins to predict their three-dimensional atomic coordinates directly from amino acid sequences. The model uses multiple sequence alignments (MSAs) and pairwise residue features as inputs, processed through a specialized network block. The subsequent structure module introduces explicit 3D transformations rotations and translations for each residue, allowing the network to generate highly accurate atomic-level protein structures. Structure prediction can be completed in GPU minutes to hours depending on protein length. By achieving near-experimental precision, AlphaFold provides valuable insights into protein folding and stability, making it a powerful tool for assessing and improving the structural robustness of biologics. Its ability to predict conformational changes and identify instability-prone regions significantly accelerates biologic design and development compared to traditional experimental methods [5].

APPLICATION OF AI IN STABILITY PREDICTION OF BIOLOGICS

1) Protein Stability Prediction Artificial intelligence has significantly improved protein stability prediction by identifying sequence determinants associated with biologic instability and aggregation. Structure-based models such as ProteinMPNN and evolutionary models like MSA Transformer enable prediction of stabilizing mutations and assessment of conformational stability. These AI-driven approaches assist in engineering biologics with improved structural integrity and reduced degradation risk [6].

2) Aggregation Prediction -Protein aggregation is a major cause of biologic instability, leading to loss of efficacy and increased immunogenicity. Machine learning and deep learning models can identify aggregation-prone regions by analysing amino acid sequences, hydrophobic interactions, and structural features. Early prediction of aggregation behaviour supports the development of safer and more stable biologic formulations [19].

3) AI-Based Formulation Optimization - Artificial intelligence is increasingly used in formulation optimization of biologics through prediction of suitable excipients and stabilization conditions. Machine learning tools such as the Excipient Prediction Software (ExPreSo) analyse protein structural properties, sequence information, and formulation characteristics to predict excipients associated with long-term stability. These AI-driven approaches reduce the time, cost, and experimental burden associated with conventional excipient screening while improving formulation efficiency and biologic stability [25].

REGULATORY ASPECTS –

Regulatory compliance is an essential aspect of biologic stability assessment, as stability data are used to establish shelf life, storage conditions, and packaging requirements. AI-based stability prediction approaches must align with established pharmaceutical regulatory frameworks and quality systems to ensure reliability and patient safety. International guidelines such as ICH Q1A (R2) for general stability testing and Q5C for biotechnological and biological products [7,10]. Additionally, Q8–Q10 emphasize stability testing, risk-based development, and pharmaceutical quality systems. Artificial intelligence models are increasingly explored as supportive tools for interpretation of stability data and early identification of stability risks [9,12]. Regulatory agencies including the FDA and EMA highlight the importance of data integrity, reproducibility, transparency, and model validation for AI-assisted pharmaceutical applications [28]. AI models should provide scientifically explainable predictions and clearly identify the factors influencing protein stability, aggregation, and degradation behaviour. Therefore, successful integration of AI into biologics development requires continuous monitoring, proper documentation, and compliance with evolving pharmaceutical regulatory standards. AI-based stability models must be supported by high-quality datasets and appropriate experimental validation to ensure consistent and scientifically reliable predictions [1].

CURRENT GAPS IN AI-BASED STABILITY PREDICTION

One of the major limitations of AI-driven stability testing is its strong dependence on large well accurate datasets. Many AI models are trained on specific datasets or molecular classes, limiting their ability to generalize across diverse biologic formulation, delivery systems, and manufacturing condition [6]. Regulatory agencies such as the U.S.Food and Drug Administration and European Medicines Agency require explainable, traceable, and scientifically justified AI systems before their implementation in stability evaluation. Early collaboration between pharmaceutical industries, AI developers and regulatory authorities is necessary for establishing acceptable validation and implementation standards. Limited availability of long-duration stability datasets restricts the development of robust predictive models for shelf-life estimation and real-time stability assessment. Successful adoption of AI in pharmaceutical stability testing requires closer alignment with regulatory expectations and clearer compliance framework [26].

FUTURE PROSPECTIVES

The application of AI-driven predictive analytics in drug stability studies is expected to expand significantly in the coming years, creating new opportunities for innovation and improved pharmaceutical development. As computational technologies continue to advance, several emerging trends are likely to reshape the future of stability assessment and formulation optimization. one important future direction is the integration of advanced AI techniques into stability modelling. While conventional machine learning and deep learning methods have already demonstrated considerable success, newer approaches such as reinforcement learning, generative adversarial networks (GANs), and graph neural networks offer additional potential for improving predictive accuracy and decision-making. Reinforcement learning may support optimization of formulation strategies and accelerated stability study designs through adaptive learning processes. Similarly, GANs could be applied for synthetic data generation, data augmentation, and simulation of

degradation pathways or impurity formation patterns [1].

Graph neural networks represent another promising advancement because they can efficiently capture complex molecular relationships and formulation interactions. By representing chemical structures and excipient interactions as graph-based data, these models may provide deeper insights into degradation mechanisms and stability behaviour. Such approaches have the potential to enable continuous prediction of product stability throughout the pharmaceutical lifecycle, thereby supporting proactive quality management, rapid decision-making, and early risk identification during both development and post-approval stages AI-based predictive models are also expected to contribute significantly to real-time release testing and continuous manufacturing systems. By correlating manufacturing process parameters with stability outcomes, predictive analytics may strengthen real-time quality assurance and reduce dependence on conventional end-product testing approaches. In addition, future progress in this field will depend heavily on the development of explainable and transparent AI systems. Regulatory acceptance of AI-driven stability tools requires models capable of providing scientifically interpretable predictions and clear justification for decision-making processes. As regulatory agencies gain more experience with AI-assisted analytical technologies, clearer guidelines and harmonized frameworks for validation and regulatory submission of AI-based stability data are expected to emerge [4].Furthermore, stronger collaboration among pharmaceutical industries, academic researchers, and regulatory authorities will be essential for establishing standardized validation strategies, best practices, and acceptable implementation frameworks for AI in pharmaceutical stability assessment Overall, the continued advancement of AI technologies is expected to transform stability testing into a more predictive, efficient, and data-driven process, ultimately improving product quality, development timelines, and patient safety[27].

CONCLUSION

The integration of Artificial Intelligence (AI) and Machine Learning (ML) into drug development in biologic drug development in modern biopharmaceutical science. These technologies have the potential to significantly transform the industry by improving efficiency, enhancing clinical trial design, accelerating drug discovery, and supporting personalized medicine. By analysing large and complex datasets, AI can help identify promising drug candidates, optimize molecular structures, and predict patient responses more accurately than traditional methods. However, despite these advantages, the adoption of AI in pharmaceuticals also brings important challenges. Issues such as data quality, model transparency, bias, ethical concerns, and patient data privacy must be carefully addressed. In addition, the effectiveness of AI systems depends heavily on proper validation, robust testing, and continuous human oversight to ensure safe and reliable outcomes. Regulatory authorities such as the FDA and EMA are actively developing adaptive frameworks to guide the responsible use of AI in biopharmaceuticals. These frameworks emphasize transparency, fairness, accountability, and the integration of real-world evidence while ensuring strong data integrity and risk-based decision-making.Overall, AI is expected to play a crucial role in reshaping the future of drug development and clinical research. With continued collaboration between industry, academia, and regulatory bodies, along with strong ethical governance and technological advancement, AI can significantly improve the speed, cost-

effectiveness, and success rate of developing new therapies, ultimately benefiting global healthcare outcomes.

REFERENCES

1. Mirakhori F, Niazi SK. Harnessing the AI/ML in Drug and Biological Products Discovery and Development: The Regulatory Perspective. Pharmaceuticals. 2025;18:47.
2. Tang J, Gong D, Li H, Li S. Artificial intelligence in biologic drug discovery: A review of methodological evolution and therapeutic applications. Acta Pharmaceutica Sinica B. 2026.
3. Tummala SR, Gurrepati N. AI-driven predictive analytics for drug stability studies. Journal of Pharma Insights and Research. Accepted 2024 Mar 29.
4. Deng J, Yang Z, Ojima I, Samaras D, Wang F. Artificial intelligence in drug discovery: applications and techniques. Briefings in Bioinformatics. 2022;23(3):1–19.
5. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021 Aug 26;596:583
6. Manning MC, Holcomb RE, Payne RW, Stillahn JM, Connolly BD, Katayama DS, Liu H, Matsuura JE, Murphy BM, Henry CS, Crommelin DJA. Stability of protein pharmaceuticals: recent advances. Pharm Res. 2024 Jul;41(7):1301–1367.
7. International Council for Harmonisation. ICH Q1A (R2): Stability testing of new drug substances and products. ICH; 2003.
8. International Council for Harmonisation. ICH Q1B: Photostability testing of new drug substances and products. ICH; 1996.
9. International Council for Harmonisation. ICH Q8 (R2): Pharmaceutical development. ICH; 2009.
10. International Council for Harmonisation. ICH Q5C: Quality of biotechnological Products: stability Testing of Biotechnological Products. ICH;1996.
11. International Council for Harmonisation. ICH Q9: Quality risk management. ICH; 2005.
12. International Council for Harmonisation. ICH Q10: Pharmaceutical quality system. ICH; 2008.
13. Shah R, Chirra M, Desai N. Application of artificial intelligence in pharmaceutical product development and quality assurance. AAPS PharmSciTech. 2020;21(6):1–12.
14. Bhattacharjee S, Dash AK. Machine learning approaches in pharmaceutical stability prediction: Opportunities and challenges. J Pharm Innov. 2022;17(2):451–463.
15. Ferreira AP, Tobyn M. Multivariate analysis in pharmaceutical stability testing. J Pharm Biomed Anal. 2015;110:264–272.
16. Montgomery DC. Introduction to Statistical Quality Control. 7th ed. New York: Wiley; 2013.
17. Aswad DW. Deamidation and isoaspartate formation in peptidesand proteins. Boca Raton: CRC Press; 1995.
18. Pogocki D, Ghezzo-Schoneich E, Schoneich C. Conformational flexibility controls proton transfer between the methioninehydroxy sulfuranyl radical and the N-terminal amino group inThr-(X)n-Met peptides. J Phys Chem B. 2001;105:1250–9.
19. Fink AL. Protein aggregation: folding aggregates, inclusion bodies and amyloid. Fold Des. 1998;3:R9–R23 295.
20. Bee, J.S.; Zhang, Y.; Finkner, S.; O'Berry, K.; Kaushal, A.; Phillippi, M.K.; DePaz, R.A.; Webber, K.; Marshall, T. Mechanistic Studies and Formulation Mitigations of Adeno-Associated Virus Capsid Rupture During Freezing and Thawing: Mechanisms

ofFreeze/Thaw Induced AAV Rupture. J. Pharm. Sci. 2022, 111, 1868–1878.

21. Ben Ali S, Bouzidi H, Chaabouni S. Machine learning-based prediction of pharmaceutical product stability. J Pharm Sci. 2021;110(5):1895–1903.
22. Patel P, Shah J. Use of machine learning for shelf-life prediction of pharmaceutical products. Asian J Pharm Sci. 2023;18(2):203–214.
23. Dara S, Dhamercherla S, Jadav SS, Babu CHM, Ahsan MJ. Machine learning in drug discovery: a review. Artif Intell Rev. 2022;55:1947–1999. Published online 2021 Aug 11.
24. Kant S, Deepika, Roy S. Artificial intelligence in drug discovery and development: transforming challenges into opportunities. Discover Pharm Sci. 2025 Jun 2. Accepted 2025 May 14.
25. Vidal-Henriquez E, Holder T, Lee NF, Pompe C, Teese MG. Machine learning driven acceleration of biopharmaceutical formulation development using Excipient Prediction Software (ExPreSo). bioRxiv. Preprint posted 2025 Feb 16.
26. Agarkar, Y., Kadam, A., Chavan, C., Jain, S., Udapurkar, P. (2026). AI Driven Stability Testing in Pharmaceutical Industry. International Journal of Medical and Pharmaceutical Sciences, 2(2), 136 145.
27. FDA. Using Artificial Intelligence and Machine Learning in the Development of Drug and Biological Products; US Food & Drug Administration: Silver Spring, MD, USA, 2023.
28. FDA. Artificial Intelligence in Drug Manufacturing; US Food & Drug Administration: Silver Spring, MD, USA, 2023.

HOW TO CITE: Ankita Gidde, AI-Driven Predictive Modelling for Stability Assessment of Biologic Therapeutics, Int. J. of Pharm. Sci., 2026, Vol 4, Issue 6, 979-987. https://doi.org/10.5281/zenodo.20532275